

Attorney's Docket No.: 041003.P033

UNITED STATES PATENT APPLICATION

FOR

A LOSSY METHOD FOR COMPRESSING PICTURES AND VIDEO

Inventor: Walter G. Bright

Prepared By:  
Columbia IP Law Group

Mail Correspondence To:  
4900 SW Meadows Rd.,  
Suite 109  
Lake Oswego, OR 97035

Direct Calls To:  
(503) 534-2800

"Express Mail" mailing label number: EL605443130US

00000000 52000000



5 65536 different colors. In addition to color images, there are black and white images (also known as monochrome, comprising one bit plane), and grayscale bitmaps (typically comprising four or eight bit planes).

10 Accordingly, a 1024 x 1024 pixel 24-bit bitmap requires three Mbytes of memory if it was to be stored in its original format. This is clearly excessive, and provided the original motivation for developing compression schemes to reduce the amount of memory required to store images.

20 color stands for the cyan, magenta, and yellow subtractive primary colors,  
plus black for contrast, typically used for printed materials. A larger value

5

10

15

20

The most commonly cited lossy compression scheme is the baseline algorithm of *JPEG* (the Joint Photographic Experts Group of the International Standards Organization (ISO)). Although JPEG defines several algorithms, the most commonly referred to as JPEG is the lossy, baseline algorithm. JPEG algorithms begin by separating the chroma (color) information from the luminance (brightness) information, to take advantage of the human eye's tolerance for lower color resolution. Next, the image is broken up into smaller rectangular tiles comprise 8x8 groups of pixels. The lossy algorithm then applies a mathematical technique to the tiles, known as *Discrete Cosine Transform* (DCT). After this transformation, individual pixels no longer exist. Rather, they are

represented by a series of patterns describing how rapidly the pixels vary. Since individual pixel identities are lost in DCT, this form of JPEG is called lossy.

A primary disadvantage with JPEG is that there may be

5 discontinuities at the boundaries of each tile. For example, while JPEG does a fairly good job of providing continuity (i.e., no rapid changes) in adjacent pixels within each tile, there is no provision for blending the pixels of the outlines of the tiles. As a result, some JPEG images, especially those that apply a high compression ratio, look like a set of tiles

10 that are reassembled together to form a composite image. Ideally, a viewer should not perceive that individual tiles are used. This problem is especially pronounced when scaling images. Accordingly, it is desired to provide an image and video compression scheme that provides similar benefits to JPEG compression, but eliminated the discontinuities that are

15 often caused by the compression scheme.

## SUMMARY OF THE INVENTION

5 The present invention provides a lossy compression scheme that addresses many of the limitations found in the prior art by enabling data corresponding to still images and video frames to be compressed and decompressed such that pixel discontinuities are substantially unnoticeable, even when images are scaled. The invention employs a recursive division and comparison scheme, wherein areas of a grid corresponding to the image data are divided into increasingly smaller triangles based on the level of detail contained within each triangle. For instance, areas with a lot of detail will comprise many small triangles, while areas with less detail will have larger triangles. Data is then stored defining each of the triangles and actual and/or predicted component values corresponding to pixels within each triangle. [In one embodiment, the data defining a triangle are data defining the vertices of the triangle.]  
10  
15 According to a first aspect of the invention, a method is provided for compressing data corresponding to still images and video frames comprising a plurality of pixels arranged in a grid. Initially, each pixel will have a one or more component values, such as a grayscale intensity for grayscale images or RGB component values for color images. The grid of pixels is divided into one or more rectangular areas, preferably by using a small number of squares. Each rectangular area is then evaluated

20

independently of one another in the following recursive manner. The area is divided into a number of triangles splitting the area. For each triangle, a set of predicted component values are determined, based on actual component values at the vertices of the triangle or actual component values on or proximate to the edges of the triangle. Preferably, the predicted component values are determined by interpolating the actual component values at the vertices or along the triangle edges. The predicted component values for the pixels within the triangle are then compared with the corresponding actual component values to see if a similarity threshold is met. In brief, this test is to determine whether the area of the still image or video frame occupied by the triangle when the image is decompressed and rendered is similar to how the same area appears prior to compressing the image, whereby the similarity threshold controls the level of detail that must be maintained. Preferably, the similarity threshold may be specified by a user, or set at a default value. Additionally, the user may specify that one or more particular areas within the image be maintained at a greater level of detail than other areas. Further, the user may also specify that certain "transparent" areas be not described.

If the similarity threshold is met, the processing of a current triangle is complete and the processing of a next triangle begins. If the similarity



threshold is not met, an optional texture mapping determination is performed. In this determination, the component values of the pixels within the triangle are evaluated to see if they form a pattern corresponding to a predefined texture map. If a match is found, the

5 matching texture map is recorded along with the vertices of the triangle. If no match is found, or if the texture mapping determination is not used, the current triangle is divided in half to make two new triangles, preferably along a line connecting the midpoint of the triangle's hypotenuse with its right angle corner. The foregoing process is then repeated for each of the

10 new triangles. As a result, image data are advantageously represented by a number of triangles with no discontinuities from one triangle to its adjacent neighbor. Further, the triangles will be subdivided such that there are many small triangles corresponding to areas in the image that have a lot of detail, while larger triangles may be used for areas that have

15 lesser detail and areas in which texture mapping is applicable.

In accord with further aspects of the invention, a system is provided for implementing the method comprising a processor and memory in which machine instructions are stored that when executed by the processor, perform the functions of the method discussed above. In

20 addition, an article of manufacture is provided having machine instructions stored thereon for implementing the method when the machine

**Bright – A Lossy Method for Compressing...**  
Express Mail No.: EL605443130US

The foregoing aspects and many of the attendant advantages of this invention will become more readily appreciated as the same becomes better understood by reference to the following detailed description, when taken in conjunction with the accompanying drawings, wherein:

FIGURE 1B is a flowchart illustrating the logic used by the present invention when compressing image data corresponding to a color image;

FIGURE 3 depicts an XYZ coordinate axis set for illustrating how image data corresponding to each pixel is maintained, wherein the X and Y axis correspond to the X and Y position of a pixel within an image, and the Z value corresponds to a grayscale or a color component value;

FIGURE 5A is an exemplary image prior to compression by the present invention:

FIGURE 5C shows the exemplary set of triangles apart from the  
5 image;

FIGURE 7 illustrates an exemplary system for implementing the present invention.

10

DETAILED DESCRIPTION

With reference to the logic flowchart of FIGURES 1A-B, the method for compressing image data in accord with the present invention starts in a block 10 in which the image area (i.e., pixilated grid for a still image or video frame) is divided into one or more rectangular elements, preferably squares. Although rectangles of any shape may be used, using squares will ensure that each triangle will comprise sides coincident with a set of pixels, and that each triangle formed will comprise a right triangle, simplifying the interpolation calculations discussed below. In addition, the squares should be configured such that a relatively small number of non-overlapping squares are used. Further, description of transparent areas may be skipped.

FIGURE 2 shows a 20 x 20 pixel grid 12 comprising four corner points 1-4 and four points within the grid labeled 5-8 corresponding to an exemplary square area created by block 20. For simplicity, an area comprising only a small number of pixels are considered in the following discussion. It will be understood by those skilled in the art that actual still images and video frames may comprise 1000 or more pixels along each axis, and that a given square or rectangular area might comprise sides with similar lengths. Each point in grid 10 is defined by three component values including an X component, a Y component, and a Z component,

5 component value of the pixel. As explained in further detail below, for grayscale images a single grayscale intensity component value will be used, while for color images, three component values will generally be used.

In a first recursive iteration, the following process is applied to each square defined above, as indicated by a start loop block 14. In accord with a block 16, each square is divided along a diagonal to form two triangles. For example, grid 12 is divided along a line 18 to form a triangle 20 having vertices at corners 1, 2 and 3, and a triangle 22 having vertices at points 2, 3, and 4. Each of triangles 20 and 22 define a plane in XYZ space, wherein line 18 comprises a shared edge between the planes, and the Z component corresponds to a “height” of the plane.

20

First, a grayscale intensity value is determined at each vertex of the triangle, as provided by a block 28. For example, if triangle 20 is currently being evaluated, grayscale intensity values will be determined for the pixels corresponding to points 1, 2, and 3. A grayscale intensity value for each pixel within the triangle is then interpolated from the vertex values in a block 30. In general, the interpolation will comprises a linear interpolation, but other interpolations may be used as well, depending on the rate of change in the grayscale intensity of the pixels across a given triangle. For example, a second order interpolation may be better suited than a first order interpolation. The order of the interpolation will depend on the actual component values of the pixels. In a block 32, a similarity calculation is then made to determine a relative difference between the actual grayscale intensity values of the pixels within the triangle and the interpolated values by taking summing the squared differences and taking the square root of the sum divided by the area of the triangle. In one embodiment, a second criterion based on the difference between the actual intensity and the interpolated intensity is also employed to determine similarity.

**Bright – A Lossy Method for Compressing...**  
Express Mail No.: EL605443130US

5  
10  
15  
20

15

20



texture patterns using one or more pattern recognition algorithms that are well-known in the art. If texture mapping is applicable, the predefined texture map pattern is identified along with intensity data in a block 36. The intensity data controls how the texture map is applied during the

5   decompression process described below.

If no match with a predefined pattern is found, the current triangle is divided to create two new triangles in a block 37. Preferably, the triangle is divided by creating a line from the midpoint of the hypotenuse of the triangle to its right angle corner. For example, triangles 20 and 22

10   were created with dividing line 18, which is the hypotenuse for both triangles 20 and 22 and has a midpoint at a point 5. Accordingly, a dividing line 38 is used to connect point 5 with point 1 (the right angle corner of triangle 20) to form two new triangles 40 and 42. In alternate embodiments, non-equal triangles may be used instead. The logic then

15   flows back to block 26 and the process is repeated again for triangle 40.

Suppose that triangle 40 contains little detail such that the similarity threshold value is not exceeded when triangle 40 is evaluated, or a suitable texture map is identified. Accordingly, the logic will proceed to a decision block 44 in which a determination is made to whether there are

20   any more triangles or squares to process. In this case, triangles 42 and 22 still have to be evaluated, so the answer to decision block 44 is yes,

This time, suppose that triangle 42 contains a moderate level of detail towards its top left-hand corner, resulting in the threshold value being exceeded. As a result, in block 36 triangle 42 is divided by a line extending from a point 6 of the hypotenuse of triangle 42 to point 5, which corresponds with the right angle corner of triangle 42, thereby creating a new pair of triangles 48 and 50 (which for ease of understanding, are equal triangles). The logic then flows back to start block 26 in which triangle 48 is evaluated. Upon evaluation of triangle 48, it is determined that it needs to be divided in half by creating a line 52 from a point 7 (the midpoint of its hypotenuse) to point 6 (its right angle corner), thereby creating triangles 54 and 56. Similarly, upon evaluation of triangle 54, it is also determined that this triangle needs to be divided in half, which is accomplished by creating a line 58 from a point 8 (at the midpoint of the hypotenuse for triangle 54 to point 7 (the right hand corner for triangle 54), creating triangles 60 and 62.

20

triangles 62, 56, 50 and 22 are still yet to be evaluated. Further suppose that when each of triangles 62, 56, 50 and 22 are evaluated, the similarity threshold value is not exceeded, resulting in no additional triangles being produced. After the evaluation of triangle 22, the answer to decision

- 5 block 44 is no to triangles (all the triangles in the square (i.e., grid 12) have been evaluated), and no to squares if no more squares are to be processed. If another square needs to be processed, the logic flows back to start block 14, and the preceding process is repeated iteratively until all of the triangles in the new square meet the similarity threshold criteria. At
- 10 this point, data defining each triangle and actual and/or predicted pixel component values will be stored in memory or on a persistent storage device as a file, preferably in a binary form. In one embodiment, the data defining a triangle are data defining the vertices of the triangle. Further, the data will indicate whether extrapolation is employed, and if employed,
- 15 the kind of extrapolation employed. For example, suppose that linear interpolation is used for all of the triangles within an image. In this case, information in a header portion of a data file could indicate that linear interpolation was used on all triangles, whereby the only information that would be needed is the XY values of the vertices of each triangle, and the
- 20 Z component value(s) for the pixels at those vertices. In other instances, the data may include information corresponding to the interpolation of

5 discontinuities from one triangle to its adjacent neighbor.

Once processing of all of the squares have been completed and the corresponding data has been stored, the logic flows to a block 64 in which an optional data compression function is performed to reduce the size of a resulting compressed image data. For instance, an initial compressed image file may comprise a list of the vertices of the triangles and corresponding grayscale intensities values that are determined during the processing of all the squares within the current image being compressed. This list typically will be in the form of a string of bits or characters. In addition, if texture mapping is applicable for any of the triangles, the compressed image data will include a string or data structure for each instance of texture mapping identifying the vertices of the triangle and the appropriate pre-defined texture map. These strings may then be compressed in block 64 by one of several well-known data compression algorithms, including Huffman encoding, LZ (Lempel and Ziv) and LZW (Lempel, Ziv and Welch) compression, and arithmetic compression.

The RGB model is used mainly in color image acquisition and display. For example, in a 24-bit display system, 8 bits is allocated for each of the red, green and blue color components. Additionally, display monitors are designed with three sets of electron guns and corresponding masks, one for each of the primary colors. As a result, it generally is necessary to start with an RGB representation of an image or video frame when compressing these images.

**Bright – A Lossy Method for Compressing...**  
Express Mail No.: EL605443130US

This has something to do with the color perception of the HVS. It is known that the HVS is more sensitive to green than to red, and is le

Allocating data in accord with the perceptions of the HVS results provides more efficient encoding.

In a luminance-chrominance color models, luminance is concerned with the perceived brightness, while chrominance is related to the perception of hue and saturation of color. Roughly speaking, the luminance-chrominance representation agrees more with the color perception of the HVS. This feature makes luminance-chrominance color models more suitable for color image processing.

Common luminance-chrominance color models that are implemented in imaging systems include the HIS (Hue, Intensity, Saturation) model, the YUV model (used for PAL (Phase Alternating Line) TV systems), the YIQ model (used for NTSC (National Television Systems Committee) TV systems), the YDbDr model (used in the SECAM (Sequential Couleur a Memoire) TV system), and the YCbCr model, which is used for the JPEG and MPEG international coding standards.

Under the YCbCr model, Y corresponds to the luminance, while Cb and Cr correspond to chrominance values. The following equation is used to convert RGB values to YCbCr values.

$$\begin{pmatrix} Y \\ Cb \\ Cr \end{pmatrix} = \begin{pmatrix} 0.257 & 0.504 & 0.098 \\ -0.148 & -0.291 & 0.439 \\ 0.439 & -0.368 & -0.071 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} + \begin{pmatrix} 16 \\ 128 \\ 128 \end{pmatrix} \quad (1)$$

5 It is well-known that the HVS is much more sensitive to the luminance component than to the chrominance components. This leads to a common practice in color image and video encoding: using full resolution for the intensity component, while using a 2 by 1 subsampling both horizontally and vertically for the two chrominance components.

10 In accord with the foregoing knowledge of the HVS, the present invention preferably implements the YCbCr model when compressing color images, wherein the "Y" (i.e., luminance) component preferably is used in determining whether the similarity threshold is met. It is noted that other luminance-chrominance models may be used, as well as the original  
15 RGB data, and that the chrominance components may also be considered in determining whether the similarity threshold is met.

With reference to FIGURE 1B, the process for compressing color images begins in a start loop block 66, which is applied to each triangle produced in a recursive fashion. First, the RGB color component values  
20 for the pixels are converted to YCbCr values per equation 1 in a block 68.

While this step is shown within the processing loop of each triangle, this step may be performed on all of the pixels within the image or video frame prior to dividing the image area into squares, or for each square.

The following steps are then applied to each color component, as

5 indicated by a start loop block 70. First, in a block 72 the value of the component is determined for each vertex point. Estimated values for each of the pixels in the triangle are then calculated (preferably by interpolation) using the component values at the vertex points, as provided by a block 74. In a block 76, a similarity comparison calculation

10 is then performed comprising summing the squared differences between the interpolated values and the actual values of the color component for each pixel and taking the square root of the sum and dividing by the area of the triangle. The calculated similarity value is then compared with a threshold value in a decision block 78 to determine whether the triangle

15 needs to be further divided. If the threshold is not exceeded and further color components need to be evaluated for the triangle, the logic loops back from a loop end block 79 to start loop block 70, and the next color component is evaluated using the same process. However, it should be noted that it is preferable to use a lower threshold value for the luminance

20 (Y) component than the chrominance (Yb, Yr) components, since the HVS is more sensitive to luminance. Furthermore, in an alternative



If the threshold is exceeded for any of the color components (if all three components are evaluated) or for the luminance component (if only the luminance component is evaluated), the logic flows to an optional decision block 80 in which it is determined whether the pixel component values correspond to a predefined texture map, in a manner similar to that described above. If texture mapping is applicable, the texture map data is saved in a block 81, and the logic proceeds to end loop block 79. If texture mapping isn't applicable, the logic flows to a block 82 that divides the triangle from the midpoint of its hypotenuse to its right angle corner, and the logic loops back to start block 66 to begin evaluation of one of the new triangles that is created.

After each of the color components has been evaluated, the logic flows to a decision block 84 in which a determination is made to whether there are any more triangles or squares to process. If the answer is yes, the logic loops back to the appropriate loop start block and processing of the next triangle or square begins. After all of the squares and triangles have been processed, data is generated comprising the X and Y values of each vertex and the color component values for that vertex. If texture

5 compressed in a block 86 using one of the data compression schemes  
discussed above.

## Pixel Interpolation Based on Edge Data

In some instances, especially when there is a large amount of detail in an image, it may be advantageous to interpolate the pixel component values based on the actual values of the pixels in the edges (i.e., lines) that form the triangles rather than using just the three vertex values. For example, consider a triangle 88 shown in FIGURE 4, which includes a plurality of pixels 90. Each pixel 90 has a grayscale intensity value disposed adjacent to it corresponding to a scale from 0 to 7, wherein 0 represents a minimum intensity level and 7 represents a maximum intensity level.

As depicted in the drawing, there are several bands having different intensity levels that run across triangle 88. For example, each of the pixels with a grayscale intensity value of 6 comprises a portion of a band, while groups of pixels with intensity value of 3 and 1 respectively comprise other bands. Bands of this type are often encountered within

images.

Under the present embodiment, the intensity values for the pixels within the boundary defined by the edges of a given triangle are interpolated based on the values of the pixels in the triangle edges. As a result, banded regions within triangles will generally be stored in the compressed image data such that these regions are accurately reproduced upon decompressing the data. Furthermore, the number of triangles that are required to create an accurate image definition may be sharply reduced, as is shown by the dashed lines in the Figure, wherein these lines represent triangles that would normally have been produced if only the vertex data was used for interpolation.

In addition to interpolation using triangle edge data, it also may be desirable to ensure that individual triangle edges are not recognized when the image data is decompressed. This can be accomplished by having the three vertices of each triangle define a surface by setting the second derivative with respect to the color of a raster line on the image as it crosses the line between two vertices to zero. This will smooth out any abrupt change in the slope, resulting in triangles that are substantially invisible to the human eye, even at the highest, lossiest compression settings.

An application of the present invention when used to compress an

exemplary grayscale image is now described with reference to FIGURES 5A-C. FIGURE 5A shows a still grayscale image 100 of a cat in a doorway, comprising 600 pixels high by 400 pixels wide. FIGURE 5B illustrates an exemplary set of triangles 102 that are created during compression of image 100 by the present invention, wherein the triangles are overlaid on top of the image. FIGURE 5C shows triangle set 102 separately.

In accord with the method described above, image 100 is first divided into a (preferably minimum) number of squares, including a large square 104 comprising 400 x 400 pixels, and two smaller squares 106 and 108, comprising 200 x 200 pixels each. Beginning with square 104, the square is first divided in half along one of its diagonals, and then the similarity comparison threshold evaluation corresponding to blocks 28-36 in FIGURE 1A is performed reiteratively, creating multiple triangles, each subdividing a previous triangular area in half. As can be recognized by examining FIGURE 5B, the density of the triangles is approximately proportional to the level of detail in each area of square 104 (and for square 106 and 108 as well). For example, small triangles are necessary to accurately store data corresponding to the details of the cat and the curves in the gate, while larger triangles may be used for lower detail areas, such as the wall on the right-hand side of the image. Smaller

triangles are also necessary for high-contrast areas, such as the near-vertical lines appearing behind the cat. In addition, where applicable, texture mapping may be performed.

The same process that was used for square 104 is then applied to  
5 each of squares 106, resulting in a set of image data in a compressed  
format. The set of image data may then be compressed using one of the  
well-known data compression algorithms discussed above.

As discussed above, the method additionally enables a user to  
define one or more areas within an image that are to be compressed such  
10 that the detail in these areas is preserved at a higher level than other  
areas in the image that are not selected. For example, the user may  
determine that he wants the cat in FIGURE 5A to be preserved with a  
higher degree of detail than other areas in the image. Accordingly, the  
user may draw a box 110 around the cat (through use of an appropriate  
15 user interface) to identify to an application program implementing the  
invention that the user wishes this area of the image to be compressed  
such that the area's detail is preserved at a higher level than other areas  
in the drawing. There are several ways this can be implemented,  
including lowering the threshold value for triangles that are contained  
20 within and/or overlap the selected area(s).

Decompressing the image data to produce an image essentially comprises reversing the steps used in creating the compressed image data discussed above. With reference to FIGURE 6, the decompression

15 In a decision block 118, a determination is made to whether texture  
mapping is to applied to the triangle. If the answer is yes, than an  
appropriate predetermined texture mapping pattern is used to create  
component intensity values for each of the pixels within the triangle, as  
provided by a block 120. If the answer is no, the interpolated values for  
20 each of the pixels in the triangle are determined in a block 122 through  
use of either the vertices data or triangle edge data and corresponding

The logic flows next to a decision block 124 in which a determination is made to whether the data corresponds to a color image. If the answer is yes, the logic reaches an end block 126 that causes the process to loop back to start block 116 and begin processing the next color component if additional color components are left to be evaluated. Once all of the color components have been processed, they are converted from luminance/chrominance values to RGB values in a block 128. Note that if the data was never converted from RGB to luminance/chrominance values this step does not occur.

After the data has been converted into an appropriate component set (e.g. RGB or grayscale), the pixels for the triangle are rendered in a block 130, whereupon the logic reaches an end block 132 that causes the logic to loop back to begin loop 114 and begin processing of the next triangle. After all of the triangles have been processed, rendering of the decompressed image is complete.

The rendering process creates varying effects, depending on the order in which the triangles are rendered. For example, the compressed data may be rearranged so that the largest enclosing triangles are rendered first, followed by the next enclosed interior triangle. As a result, the image will start out appearing fuzzy and will get progressively sharper.

5

10

15



by the personal computer, and for graphically representing images and video frames produced by the present invention. A mouse 210 (or other pointing device) is connected to a serial port (or to a bus port) on the rear of processor chassis 202, and signals from mouse 210 are conveyed to

5 the motherboard to control a cursor on the display and to select text, menu options, and graphic components displayed on monitor 208 by software programs executing on the personal computer, such as a photo editing program that implements the present invention. In addition, a keyboard 212 is coupled to the motherboard for user entry of text and

10 commands that affect the running of software programs executing on the personal computer.

Personal computer 200 also optionally includes a compact disk-read only memory (CD-ROM) drive 214 into which a CD-ROM disk may be inserted so that executable files and data on the disk can be read

15 for transfer into the memory and/or into storage on hard drive 206 of personal computer 200. Other mass memory storage devices such as an optical recorded medium or DVD drive may be included. The machine instructions comprising the software program that causes the CPU to implement the functions of the present invention that have been

20 discussed above will likely be distributed on floppy disks or CD-ROMs (or

The above description of illustrated embodiments of the invention is not intended to be exhaustive or to limit the invention to the precise forms disclosed. While specific embodiments of, and examples for, the invention are described herein for illustrative purposes, various equivalent modifications are possible within the scope of the invention, as those skilled in the relevant art will recognize. Accordingly, it is not intended that the scope of the invention in any way be limited by the above description, but instead be determined entirely by reference to the claims that follow.